Computers, Environment and Urban Systems 65 (2017) 126–139

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks

Geoff Boeing

Department of City and Regional Planning, University of California, Berkeley, 228 Wurster Hall #1850, Berkeley, CA 94720-1850, United States

ARTICLE INFO

Article history:
Received 28 November 2016
Received in revised form 22 April 2017
Accepted 26 May 2017
Available online 7 July 2017

Keywords:
Street network
Urban form
Urban design
Transportation
Resilience
OpenStreetMap
GIS
Complex networks
Visualization
Python

ABSTRACT

Urban scholars have studied street networks in various ways, but there are data availability and consistency limitations to the current urban planning/street network analysis literature. To address these challenges, this article presents OSMnx, a new tool to make the collection of data and creation and analysis of street networks simple, consistent, automatable and sound from the perspectives of graph theory, transportation, and urban design. OSMnx contributes five significant capabilities for researchers and practitioners: first, the automated downloading of political boundaries and building footprints; second, the tailored and automated downloading and constructing of street network data from OpenStreetMap; third, the algorithmic correction of network topology; fourth, the ability to save street networks to disk as shapefiles, GraphML, or SVG files; and fifth, the ability to analyze street networks, including calculating routes, projecting and visualizing networks, and calculating metric and topological measures. These measures include those common in urban design and transportation studies, as well as advanced measures of the structure and topology of the network. Finally, this article presents a simple case study using OSMnx to construct and analyze street networks in Portland, Oregon.

1. Introduction

Urban scholars and planners have studied street networks in numerous ways. Some studies focus on the urban form (e.g., Southworth & Ben-Joseph, 1997; Strano et al., 2013), others on transportation (e.g., Marshall & Garrick, 2010; Parthasarathi, Levinson, & Hochmair, 2013), and others on the topology, complexity, and resilience of street networks (e.g., Jiang & Claramunt, 2004; Porta, Crucitti, & Latora, 2006). This article argues that current limitations of data availability, consistency, and technology have made researchers' work gratuitously difficult. In turn, this empirical literature often suffers from four shortcomings which this article examines: small sample sizes, excessive network simplification, difficult reproducibility, and the lack of consistent, easy-to-use research tools. These shortcomings are by no means fatal, but their presence limits the scalability, generalizability, and interpretability of empirical street network research.

To address these challenges, this article presents OSMnx, a new tool that easily downloads and analyzes street networks for anywhere in the world. OSMnx contributes five primary capabilities for researchers and practitioners. First, it enables automated and on-demand downloading of political boundary geometries, building footprints, and elevations. Second, it can automate and customize the downloading of street networks from OpenStreetMap and construct them into multidigraphs. Third, it can correct and simplify network topology. Fourth, it can save/load street networks to/from disk in various file formats. Fifth and finally, OSMnx has built-in functions to analyze street networks, calculate routes, project and visualize networks, and quickly and consistently calculate various metric and topological measures. These measures include those common in urban design and transportation studies, as well as advanced measures of the structure and topology of the network.

This article is organized as follows. First, it introduces the background of networks, street network analysis and representation, and the current landscape of tools for this type of research. Then it discusses shortcomings and current challenges, situated in the empirical literature. Next, it introduces OSMnx and its methodological contributions. Finally, it presents a simple illustrative case study using OSMnx to construct and analyze street networks in Portland, Oregon, before concluding with a discussion.

2. Background

Street network analysis has been central to network science since its nascence: its mathematical foundation, graph theory, was born in the 18th century when Leonhard Euler presented his famous Seven Bridges of Königsberg problem. Here we briefly trace the fundamentals of

E-mail address: gboeing@berkeley.edu.

http://dx.doi.org/10.1016/j.compenvurbsys.2017.05.004
0198-9715/

modern street network research from graphs to networks to the present landscape of research toolkits, in order to identify current limitations.

2.1. Graphs and networks

Network science is built upon the foundation of graph theory, a branch of discrete mathematics. A *graph* is an abstract representation of a set of elements and the connections between them (Trudeau, 1994). The elements are interchangeably called vertices or nodes, and the connections between them are called links or edges. For consistency, this article uses the terms *nodes* and *edges*. The number of nodes in the graph (called the *degree* of the graph) is commonly represented as n and the number of edges as m. Two nodes are *adjacent* if an edge connects them, two edges are adjacent if they share the same node, and a node and an edge are *incident* if the edge connects the node to another node. A node's *degree* is the number of edges incident to the node, and its *neighbors* are all those nodes to which the node is connected by edges.

An *undirected* graph's edges point mutually in both directions, but a *directed* graph, or digraph, has directed edges (i.e., edge uv points from node u to node v, but there is not necessarily a reciprocal edge vu). A *self-loop* is an edge that connects a single node to itself. Graphs can also have parallel (i.e., multiple) edges between the same two nodes. Such graphs are called *multigraphs*, or *multidigraphs* if they are directed.

An undirected graph is *connected* if each of its nodes can be reached from any other node. A digraph is *weakly connected* if the undirected representation of the graph is connected, and *strongly connected* if each of its nodes can be reached from any other node. A *path* is an ordered sequence of edges that connects some ordered sequence of nodes. Two paths are *internally node-disjoint* if they have no nodes in common, besides end points. A *weighted* graph's edges have a weight attribute to quantify some value, such as importance or impedance, between connected nodes. The *distance* between two nodes is the number of edges in the path between them, while the *weighted distance* is the sum of the weight attributes of the edges in the path.

While a graph is an abstract mathematical representation of elements and their connections, a *network* may be thought of as a real-world graph. Networks inherit the terminology of graph theory. Familiar examples include social networks (where the nodes are humans and the edges are their interpersonal relationships) and the World Wide Web (where the nodes are web pages and the edges are hyperlinks that point from one to another). A *complex* network is one with a nontrivial *topology* (the configuration and structure of its nodes and edges) – that is, the topology is neither fully regular nor fully random. Most large real-world networks are complex (Newman, 2010). Of particular interest to this study are *complex spatial networks* – that is, complex networks with nodes and/or edges embedded in space (O'Sullivan, 2014). A street network is an example of a complex spatial network with both nodes and edges embedded in space, as are railways, power grids, and water and sewage networks (Barthélemy, 2011).

2.2. Representation of street networks

A spatial network is *planar* if it can be represented in two dimensions with its edges intersecting only at nodes. A street network, for instance, *may* be planar (particularly at certain small scales), but most street networks are non-planar due to grade-separated expressways, overpasses, bridges, and tunnels. Despite this, most quantitative studies of urban street networks represent them as planar (e.g., Barthélemy & Flammini, 2008; Buhl et al., 2006; Cardillo, Scellato, Latora, & Porta, 2006; Masucci, Smith, Crooks, & Batty, 2009; Strano et al., 2013) for tractability because bridges and tunnels are reasonably uncommon (in certain places) – thus the networks are *approximately* planar. However, this over-simplification to planarity for tractability may be unnecessary and can cause analytical problems, as we discuss shortly.

The street networks discussed so far are *primal*: the graphs represent intersections as nodes and street segments as edges. In contrast, a *dual graph* (namely, the edge-to-node dual graph, also called the *line graph*) inverts this topology: it represents a city's streets as nodes and intersections as edges (Porta et al., 2006). Such a representation may seem a bit odd but provides certain advantages in analyzing the network topology based on named streets (Crucitti, Latora, & Porta, 2006). Dual graphs form the foundation of space syntax, a method of analyzing urban networks and configuration via axial street lines and the depth from one edge to others (Hillier, Leaman, Stansall, & Bedford, 1976; cf. Ratti, 2004). Jiang and Claramunt (2002) integrate an adapted space syntax – compensating for difficulties with axial lines – into computational GIS. Space syntax has formed the basis of various other adapted approaches to analytical urban design (e.g., Karimi, 2012).

This present article, however, focuses on primal graphs because they retain all the geographic, spatial, metric information essential to urban form and design that dual representations discard: all the geographic, experiential traits of the street (such as its length, shape, circuity, width, etc.) are lost in a dual graph. A primal graph, by contrast, can faithfully represent all the spatial characteristics of a street. Primal may be a better approach for analyzing spatial networks when geography matters, because the physical space underlying the network contains relevant information that cannot exist in the network's topology alone (Ratti, 2004).

2.3. Street network analysis

Street networks – considered here as primal, non-planar, weighted multidigraphs with self-loops – can be characterized and described by metric and topological measures. Extended definitions and algorithms can be found in, e.g., Newman (2010) and Barthélemy (2011).

Metric structure can be measured in terms of length and area and represents common transportation/design variables (e.g., Cervero & Kockelman, 1997; Ewing & Cervero, 2010). *Average street length*, the mean edge length (in spatial units such as meters) in the undirected representation of the graph, serves as a linear proxy for block size and indicates how fine-grained or coarse-grained the network is. *Node density* is the number of nodes divided by the area covered by the network. *Intersection density* is the node density of the set of nodes with more than one street emanating from them (thus excluding dead-ends). The *edge density* is the sum of all edge lengths divided by the area, and the physical *street density* is the sum of all edges in the undirected representation of the graph divided by the area. These density measures all provide further indication of how fine-grained the network is. Finally, the *average circuity* divides the sum of all edge lengths by the sum of the great-circle distances between the nodes incident to each edge (cf. Giacomin & Levinson, 2015). This is the average ratio between an edge length and the straight-line distance between the two nodes it links.

The *eccentricity* of a node is the maximum of the shortest-path weighted distances between it and each other node in the network. This represents how far the node is from the node that is furthest from it. The *diameter* of a network is the maximum eccentricity of any node in the network and the *radius* of a network is the minimum eccentricity of any node in the network. The *center* of a network is the node or set of nodes whose eccentricity equals the radius and the *periphery* of a network is the node or set of nodes whose eccentricity equals the diameter. When weighted by length, these distances indicate network size and shape in units such as meters.

Topological measures of street network structure indicate the configuration, connectedness, and robustness of the network – and how these characteristics are distributed. The *average node degree*, or mean number of edges incident to each node, quantifies how well the nodes are connected on average. Similarly, but more concretely, the *average streets per node* measures the mean number of physical streets (i.e., edges in the undirected representation of the graph) that emanate from each intersection and dead-end. This adapts the average node degree for